ENTROPY INCREASE FOR A CLASS OF DYNAMICAL MAPS

G. ANANTHAKRISHNA
Department of Physics, Indian Institute of Science, Bangalore, India

E. C. G. SUDARSHAN
Department of Physics, Center for Particle Theory
The University of Texas, Austin, USA
Centre for Theoretical Studies, Indian Institute of Science, Bangalore, India

VITTORIO GORINI
Istituto di Fisica dell' Università, Milano, Italy
Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy
(Received April 1, 1974)

The dynamical evolution of a quantum system is described by a one parameter family of linear transformations of the space of self-adjoint trace class operators (on the Hilbert space of the system) into itself, which map statistical operators to statistical operators. We call such transformations dynamical maps. We give a sufficient condition for a dynamical map $A$ not to decrease the entropy of a statistical operator. In the special case of an $N$-level system, this condition is also necessary and it is equivalent to the property that $A$ preserves the central state.

1. Introduction

In the conventional framework of quantum mechanics to every physical system $\Sigma$ is associated a complex Hilbert space $\mathcal{H}$ and a state of $\Sigma$ is represented by a positive trace class operator $\rho$ on $\mathcal{H}$ with unit trace. One also refers to $\rho$ as a statistical operator of the system. A (bounded) observable of $\Sigma$ is represented by a self-adjoint bounded operator $a$ on $\mathcal{H}$ and its expectation value in the state $\rho$ is given by $\rho(a) = \text{Tr}(\rho a)$. The most general dynamical development of the system, which will not necessarily be assumed to be isolated (and no physical system is indeed isolated in the true sense!), is provided by a one-parameter family $t \mapsto A_t$, $t \geq 0$, $A_0 = 1$, of transformations which map states to states in such a way that if $\rho_0$ represents the state at $t = 0$, $\rho_t = A_t \rho_0$ gives the state at a later time $t$. We assume on physical grounds that $(A_t \rho_0)(a) = \text{Tr}[(A_t \rho_0) a]$ is a continuous function of $t$ for every observable $a$ and for every initial state $\rho_0$. In most cases, when the surroundings are so large as not to be appreciably affected by the system in question, one can assume $A_t$ to preserve the incoherent superposition of states:

$$A_t(\alpha \rho + (1 - \alpha) \rho') = \alpha A_t \rho + (1 - \alpha) A_t \rho', \quad 0 \leq \alpha \leq 1.$$
The above framework can be formalized as follows. Let $T(\mathcal{S})$ (respectively $B(\mathcal{S})$) denote the real Banach space of the self-adjoint trace class (respectively self-adjoint bounded) operators on $\mathcal{S}$ with the trace norm $\|\sigma\|_1 = \text{Tr} |\sigma|$ (respectively with the operator norm $\|\sigma\| = \sup \|\sigma x\|$). Via the correspondence $\sigma \to \text{Tr}(\sigma a)$, $B(\mathcal{S})$ can be identified to the strong dual of $T(\mathcal{S})$ [16]. $K(\mathcal{S}) = \{ \sigma; \sigma \in T(\mathcal{S}), \sigma \geq 0, \text{Tr}\sigma = 1\}$ denotes the convex set of statistical operators (states of the system $\Sigma$). It is a base of the positive cone $T^+(\mathcal{S}) = \{ \sigma; \sigma \in T(\mathcal{S}), \sigma \geq 0\}$.

**Definition 1.1** ([18], [9]). A dynamical map is an affine map $A : K(\mathcal{S}) \to K(\mathcal{S})$, namely $A(\lambda \sigma + (1 - \lambda) \rho) = \lambda A\sigma + (1 - \lambda) A\rho$, $\sigma, \rho, A\sigma, A\rho \in K(\mathcal{S})$, $0 \leq \lambda \leq 1$.

Using the fact that a self-adjoint trace class operator $\sigma$ admits of a unique decomposition $\sigma = \sigma^+ - \sigma^-$ with $\sigma^\pm \in T^+(\mathcal{S})$, a dynamical map $A$ can be uniquely extended to a linear map of $T(\mathcal{S})$ into itself, which we again denote by $A$, by defining $A\sigma = (\text{Tr}\sigma^+) A\sigma^+ - (\text{Tr}\sigma^-) A\sigma^-$ where $\tilde{\sigma}^\pm = \sigma^\pm/\text{Tr}\sigma^\pm$ or $\tilde{\sigma}^\pm = 0$ according to whether $\sigma^\pm \neq 0$ or $\sigma^\pm = 0$, and $\sigma^0 = 0$. Therefore, we may henceforth identify the set of dynamical maps to the convex set $F(\mathcal{S})$ of linear maps of $T(\mathcal{S})$ into itself which leave $K(\mathcal{S})$ invariant.

If $A$ is a dynamical map, its dual $A^*$ is the positive and identity preserving linear map $B(\mathcal{S}) \to B(\mathcal{S})$ defined by $\text{Tr}[\sigma(A^* a)] = \text{Tr}[(A\sigma) a]$, $\sigma \in T(\mathcal{S}), a \in B(\mathcal{S})$.

**Proposition 1.1** [11]. Let $A$ be a dynamical map. Then $A$ is bounded and has norm 1.

**Proof.** $\|A\sigma\|_1 = \|A\sigma^+ - A\sigma^-\|_1 \leq \|A\sigma^+\|_1 + \|A\sigma^-\|_1 = \text{Tr} A\sigma^+ + \text{Tr} A\sigma^- = \text{Tr} \sigma^+ + \text{Tr} \sigma^- = \text{Tr} |\sigma| = \|\sigma\|_1$ and, if $\rho \in K(\mathcal{S})$, $\|A\rho\|_1 = \|\rho\|_1$. \(\blacksquare\)

**Definition 1.2.** A dynamical evolution is a one-parameter family $\mathcal{A} : t \to A_t, t \geq 0$, of dynamical maps such that a) $A_0 = 1$ and b) $\mathcal{A}$ is weakly continuous, in the sense that the function $t \to \text{Tr}(A_t a)$ is continuous $\forall a \in K(\mathcal{S})$ and $\forall a \in B(\mathcal{S})$.

To every dynamical evolution $t \to A_t$ we associate its dual $t \to A_t^*$. According to the conventional terminology we refer to $t \to A_t$ as the Schrödinger evolution and to $t \to A_t^*$ as the corresponding Heisenberg evolution. If the dynamics is invariant under time translation one has further c) $A_t A_s = A_{t+s}$ and $\mathcal{A}$ is a strongly continuous one parameter contraction semigroup (compare [19] and Proposition 1.1). It will be called a dynamical semigroup [11].

Dynamical semigroups have been studied extensively by Kossakowski [11]-[14] as regards their general structure and in connection with the description of the dynamics of certain open systems (laser theory, Bloch equations). In a similar setting, Davies [2]-[5] has studied certain one parameter strongly continuous semigroups of positive endomorphisms of $T(\mathcal{S})$, in connection with quantum stochastic processes originating from the interaction of the system with the measuring devices. If a dynamical semigroup $\mathcal{A} : t \to A_t$ describes a reversible dynamics in the sense that, for every $t \in [0, \infty)$, $A_t$ maps $K(\mathcal{S})$ onto itself one to one, then $A_t$ has the form ([17], [10], [8], [11]) $A_t = e_t e^* e_t^*$, where $t \mapsto e_t = \exp(-iHt), t \in (-\infty, \infty)$, is a strongly continuous one parameter group of unitary operators on $\mathcal{S}$ whose generator $H$ can be interpreted as the Hamiltonian of the
system. Such a Hamiltonian evolution can possibly be ascribed to the idealized situation of a strictly isolated system. On the other hand, for systems undergoing truly irreversible processes, in which an essential part is played by the participation of the incoherent surroundings, the dynamical evolution is to a greater or smaller extent non-Hamiltonian. In this connection, the question of the behaviour of the entropy of a state \( g \), \( S(g) = -\text{Tr}(g \log g) \), under a dynamical semigroup and more generally under arbitrary dynamical maps is interesting. Note that the entropy is invariant under Hamiltonian evolution unless its definition is changed, for example by the use of some coarse graining device. In this note we establish a sufficient condition for a dynamical map \( A \) to have the property that \( S(Ag) \geq S(g) \) for every \( g \in K(\mathcal{B}) \). If \( \mathcal{B} \) is finite-dimensional, the condition we give is also necessary. We suspect it to be necessary in general, but have not been able to prove it so far. Our result establishes rigorously and generalizes some previous results by Okubo and Isihara on various mechanisms of entropy increase [15].

2. Entropy increase

We define \( \hat{F}(\mathcal{B}) = \{ A \mid A \in F(\mathcal{B}); \, a \in T^{+}(\mathcal{B}) \Rightarrow \text{Tr}a \geq \text{Tr}(A^{\dagger}a) \} \). We denote by \( p_{x} \) the orthogonal projection onto the subspace spanned by the normalized vector \( x \in \mathcal{B} \). If \( \mathcal{B} = \mathbb{C}^{N} \), we call \( \frac{1}{N} I \) the central statistical operator or the centre. Our result is expressed by the following

**Proposition 2.1.** Let \( A \) be a dynamical map. In order that

\[
S(Ag) \geq S(g) \, \forall g \in K(\mathcal{B}),
\]

it is sufficient that \( A \in \hat{F}(\mathcal{B}) \). If \( \mathcal{B} = \mathbb{C}^{N} \), this condition is also necessary for (2.1) to hold and it is equivalent to the condition that \( A \) preserves the centre.

The meaning of Proposition 2.1 can be explained as follows. If \( t \rightarrow A_{t} \) is a dynamical semigroup such that \( A_{t} \in \hat{F}(\mathcal{B}) \, \forall t \geq 0 \), the condition \( \text{Tr}a \geq \text{Tr}(A_{t}^{\dagger}a) \, \forall t \geq 0 \) and \( \forall a \in T^{+}(\mathcal{B}) \) represents a tendency of the eigenvalues of the Heisenberg observable \( a = A_{t}a \) to concentrate around a common value and thus expresses a tendency towards a dispersion of information which is rightly represented by relation (2.1). From the physical point of view, if \( \mathcal{B} \) is infinite-dimensional, there exist in general constraints (such as for instance fixed values of the intensive variables of a reservoir) which prevent the system from increasing its entropy indefinitely and thus make situations such as the above unlikely, even though there appears to be no reason for them to be excluded in principle. On the other hand, Proposition 2.1 is particularly significant in the case of an \( N \)-level system, in view of the occurrence of relaxation processes ending up with the centre as equilibrium state. It is indeed a consequence of Proposition 2.1 that for the centre to be an invariant
state of a dynamical semigroup of an N-level system it is necessary and sufficient that
the entropy of the state of the system be a non-decreasing function of time, independently
of the initial conditions.

In order to prove the proposition we make use of two lemmas.

**Lemma 2.1.** Let $A$ be a dynamical map and let $(x_i)_{i=1}^\infty$ and $(y_j)_{j=1}^\infty$ be two
complete orthonormal sets in $\mathcal{S}$. Set $p_i = p_{x_i}, q_i = p_{y_i} (i, j = 1, 2, \ldots)$ and
$A_U = \text{Tr}[q_i(A p_i)]$. Then: a) $A_U$ is a stochastic matrix, namely
\begin{equation}
A_U = 0 \quad (i, j = 1, 2, \ldots),
\end{equation}
\begin{equation}
\sum_i A_{ij} = 1 \quad (j = 1, 2, \ldots);
\end{equation}
b) in order that $A \in \hat{F}({\mathcal{S}})$ it is necessary and sufficient that
\begin{equation}
\sum_j A_{ij} \leq 1 \quad (i = 1, 2, \ldots);
\end{equation}
c) if $\mathcal{S} = C^\infty$, (2.4) implies
\begin{equation}
\sum_{j=1}^N A_{ij} = 1 \quad (i = 1, 2, \ldots, N)
\end{equation}
and (2.5) is equivalent to the condition that $A$ be centre preserving.

**Proof:** a) Showing (2.2) and (2.3) is trivial. b) If $A \in \hat{F}({\mathcal{S}})$ we have $1 = \text{Tr}q_1$
$\geq \text{Tr}(A^* q_1) = \sum_j \text{Tr}[p_j(A^* q_1)] = \sum_j \text{Tr}[(A p_j) q_1] = \sum_j A_{ij} (i = 1, 2, \ldots)$. If $A \notin \hat{F}({\mathcal{S}})$ we
have $\text{Tr}(A^* a) > \text{Tr} a$ for some positive trace class operator $a$. Without loss of generality
we can assume $a$ to be a statistical operator $q$. Choose $(q_i)$ to be the spectral family of $q,$
$q = \sum_i q_i$, and choose $M$ such that $\sum_{j=1}^M (x_j, (A^* q_i) x_j) > 1.$ By the continuity of $A^*$
we have
\begin{align*}
1 = \text{Tr} q & = \sum_i \zeta_i < \sum_{j=1}^M (x_j, (A^* q_i) x_j) = \sum_{j=1}^M (x_j, [\sum_i \zeta_i (A^* q_i)] x_j) \\
& = \sum_{j=1}^M \sum_i \zeta_i (x_j, (A^* q_i) x_j) = \sum_{j=1}^M \sum_i \zeta_i \text{Tr}[p_j(A^* q_i)] \\
& = \sum_{j=1}^M \sum_i \zeta_i A_{ij} = \sum_i \zeta_i \sum_{j=1}^M A_{ij}.
\end{align*}
Hence, since $\sum \zeta_i = 1$, we have $\sum_j A_{ij} > 1$ (the l.h.s. being possibly divergent) for at
least one value of $i$. c) Let $\mathcal{S} = C^\infty$ and assume $A \in \hat{F}({\mathcal{S}})$. Then, if $x \in \mathcal{S}, \|x\| = 1$, we
have $1 = \text{Tr} p_x \geq \text{Tr} (A^* p_x) = N \text{Tr} \left[ \frac{1}{N} I (A^* p_x) \right] = N \left( x, \left[ A \left( \frac{1}{N} I \right) \right] x \right)$ or $\left( x, \left[ A \left( \frac{1}{N} I \right) \right] x \right) \leq \frac{1}{N}$ for all normalized $x \in \mathcal{S}$. Since $A \left( \frac{1}{N} I \right)$ is a statistical operator, this implies $A \left( \frac{1}{N} I \right) = \frac{1}{N} I$. Then $\sum_{j=1}^{N} A_{ij} = \sum_{j=1}^{N} \text{Tr}[q_j(A p_j)] = \text{Tr}[q_j(A I)] = \text{Tr} q_i = 1 (i = 1, 2, ..., N)$.

Conversely, from (2.5) we get $\sum_j \text{Tr}[q_i(A p_j)] = (y_i, (A I) y_i) (i = 1, 2, ..., N)$ which, by the arbitrariness of $(y_i)$, implies $A \frac{1}{N} I = \frac{1}{N} I$. $
$
**Remark 2.1.** If $\mathcal{S} = C^*$, $A \in F(\mathcal{S}) \iff A \left( \frac{1}{N} I \right) = \frac{1}{N} I \iff A^*$ is trace preserving. On the other hand, if $\mathcal{S}$ is infinite dimensional, it is possible for some $A \in F(\mathcal{S})$ that $\sum_j A_{ij} < 1$ for some value of the index $i$. For example, let $u$ be a non-unitary isometric operator on $\mathcal{S}$, $u^* u = I$, $u u^* = p = I$. It is easy to see that $A: \sigma \rightarrow uu^* \sigma, \sigma \in \mathcal{T}(\mathcal{S})$, is a dynamical map. If $\sigma$ is a positive trace class operator we have $\text{Tr} \sigma = \text{Tr} (I-p) \sigma + \text{Tr} u^* \sigma = \text{Tr} u^* uu = \text{Tr} (A^* \sigma)$. Thus $A \in F(\mathcal{S})$. Choose $(y_i)$ such that $py_i = 0$ for some $i = i_0$. Then

$$\sum_j A_{ij} = \sum_j \text{Tr}(q_j u^* q_i u) = \sum_j \text{Tr}(u^* q_i u)$$

$$= \sum_j (ux_j, q_i u x_j) = \sum_j |(ux_j, u y_i)|^2 = 0.$$

**Lemma 2.2.** Let $A \in F(\mathcal{S})$, $\sigma \in K(\mathcal{S})$, and let $(p_i)_{i=1,2,...}$ and $(q_i)_{i=1,2,...}$ be the spectral families of $\sigma$ and $A \sigma$ respectively: $\sigma = \sum_i p_i$ and $A \sigma = \sum_i (A \sigma) q_i$. Then

$$(A \sigma)_i = \sum_j \text{Tr}[q_j(A p_j)] q_i = \sum_j A_{ij} q_i \quad (i = 1, 2, ...). \quad (2.6)$$

**Proof:** Set $q_{i,n} = \sum_{i=1}^{n} q_i p_i$, $(n = 1, 2, ...)$ and let $x \in \mathcal{S}$. Then $(A \sigma) x \rightarrow (A \sigma) x$ in norm. Indeed, we have $\| (A \sigma) x - (A \sigma) x \| \leq \| A \sigma - A \sigma \| x \| x \| = \| A \sigma - A \sigma \| x \| x \| = \| x \| \sum_{n=1}^{\infty} q_i \rightarrow 0$. Hence

$$(A \sigma)_i = \text{Tr}[q_i(A \sigma)] = (y_i, (A \sigma) y_i) = \sum_j (y_i, (A p_j) y_i) q_i = \sum_j \text{Tr}[q_j(A p_j)] q_i.$$

**Proof of the Proposition:** Let $A \in F(\mathcal{S})$. Using (2.2), (2.3), (2.4), (2.6) and the convexity of the function $\sigma \rightarrow \text{Tr} \sigma$ ($\sigma > 0$) and setting $\sum_j A_{ij} = \beta_i < 1 = \sum_j B_{ij}$ ($i = 1, 2, ...$).
with $B_{ij} \geq 0$ and $B_{ij} = A_{ij}/\beta_i$ if $\beta_i \neq 0$, we get (obviously, we need only consider the case when $S(\alpha_0) < \infty$)

$$S(\alpha_0) = -\text{Tr}[(A_\alpha) \log(A_\alpha)] = -\sum_i (A_\alpha)_i \log(A_\alpha)_i$$

$$= -\sum_i \left( \sum_j A_{ij}(\log B_{ij}) \right)$$

$$= -\sum_i \left( \sum_j A_{ij}(\log B_{ij}) - \sum_j A_{ij}(\log B_{ij}) \right)$$

$$\geq -\sum_i \log B_{ij}$$

$$\geq -\sum_i \log B_{ij}$$

$$= -\sum_i \left( \sum_j A_{ij} \log B_{ij} \right)$$

$$= -\sum_i \left( \sum_j A_{ij} \log B_{ij} \right)$$

$$= -S(\alpha_0).$$

This also shows that if $S(\alpha_0) < \infty$, the equality holds in (2.1) iff $A_\alpha$ and $\alpha$ have the same eigenvalues with the same multiplicities.

If $S = C^N$ the function $\alpha \to S(\alpha)$ has a strict absolute maximum at $1/N$. Hence (2.1) implies $A$ to be centre preserving. The equivalence of the condition $A \in F(S)$ to the condition that $A$ be centre preserving was proved in Lemma 2.1.

Remark 2.2 If $S = C^N$, we can give an alternative elegant proof that $A \left( \frac{1}{N} I \right) = \frac{1}{N} I$ implies (2.1) using the concavity of the entropy function and Birkhoff's theorem on doubly stochastic $N \times N$ matrices. An $N \times N$ matrix is called doubly stochastic if it satisfies (2.2), (2.3) and (2.5) and Birkhoff's theorem asserts that the extreme elements of the convex set of the $N \times N$ doubly stochastic matrices are the $N!$ permutation matrices $P^T = \delta_{uv}$, where $u = (v_1, v_2, ..., v_N)$ is a permutation of $(1, 2, ..., N)$. With the notations of Lemma 2.2 and representing $\beta$ by the probability vector $\{p_i\}$ and $A \in F(S)$ by the doubly stochastic matrix $A_{ij}$, we have

$$A = \sum_{(v)} \lambda_{(v)} B^{(v)}, \quad \lambda_{(v)} \geq 0, \quad \sum_{(v)} \lambda_{(v)} = 1$$

and

$$S(\alpha_0) = S\left( \sum_{(v)} \lambda_{(v)} B^{(v)} \right) \geq \sum_{(v)} \lambda_{(v)} S(B^{(v)}) = S(\alpha_0).$$

\footnote{See [1]. An elementary arithmetical proof of this theorem has been recently given by N. Mukunda (private communication).}
Remark 2.3. The question arises whether (2.1) implies that \( \mathcal{A} \in \hat{F}(\mathcal{S}) \) also when \( \mathcal{S} \) is infinite dimensional. If \( \mathcal{A} \notin \hat{F}(\mathcal{S}) \) it follows from Lemma 2.1 that we can choose complete orthogonal sets of one dimensional projections \( \{p_j\}_{j=1,2,\ldots} \) and \( \{q_i\}_{i=1,2,\ldots} \) such that \( \sum_j \text{Tr}(q_i(A)p_j) = \sum_j A_{i,j} > 1 \) for some \( i = i_0 \). Without loss of generality we can choose \( i_0 = 1 \). If \( \{\pi_i^{(N)}\}_{i=1,2,\ldots} \) denotes the probability vector whose first \( N \) components equal \( 1/N \) and we write \( \pi_i^{(N)} = \sum_j A_{i,j} \pi_j^{(N)} \), for \( N \) large enough \( \pi_i^{(N)} \) becomes larger than \( 1/N \).

Then, it might be possible that for a suitable choice of \( N \) (depending on \( \mathcal{A} \)) we could get \( S(\omega^{(N)}) < \lg N \), which would answer positively to the question above.

Remark 2.4. Note that if \( \mathcal{S} = C^\pi \), \( \mathcal{K}(\mathcal{S}) \) is compact and hence any dynamical map has at least one fixed point. On the other hand, if \( \mathcal{S} \) is infinite-dimensional, a dynamical map need not have a fixed point. This applies in particular to the elements of \( \hat{F}(\mathcal{S}) \) (there is no such thing as a central element in this case!). For example, if \( \omega \) is a unitary operator whose spectrum is purely continuous, the dynamical map \( \sigma \to \omega \omega^* \) has no fixed points.

Remark 2.5. \( \hat{F}(\mathcal{S}) \) is a convex set. It follows from a theorem of Davies (see [4, Theorem 3.1]) that if \( \mathcal{A} \) belongs to \( \hat{F}(\mathcal{S}) \) and it maps pure states to pure states, then either

\[
A : \sigma \to \omega \sigma \omega^*, \quad \sigma \in \mathcal{I}(\mathcal{S}),
\]

(2.7)

where \( \omega \) is an isometric operator on \( \mathcal{S} \), or

\[
A : \sigma \to \omega \sigma \omega^*, \quad \sigma \in \mathcal{I}(\mathcal{S}),
\]

(2.8)

where \( \sigma \) is an antilinear isometric operator on \( \mathcal{S} \). Because dynamical maps are continuous, both (2.7) and (2.8) are extreme elements of \( F(\mathcal{S}) \) (and hence of \( \hat{F}(\mathcal{S}) \)) since they map pure states to pure states. Setting \( \hat{F}(\mathcal{S}) = \{A | A \in \hat{F}(\mathcal{S}); q \text{ a pure state } \Rightarrow AQ \text{ a pure state}\} \), it is an interesting question whether \( \hat{F}(\mathcal{S}) \) is the uniform closure of the convex hull of \( \hat{F}(\mathcal{S}) \). So far, to our knowledge, it is only known that this result is true in the simplest case \( \mathcal{S} = C^2 \) [6].

Acknowledgments

One of us, V.G., acknowledges a discussion with C. Cerrignani which led to a simplification of the proof of Proposition 2.1. The second author, E.C.G.S., is thankful to N. Mukunda for discussions.

REFERENCES